



TITLE:

量子アニーリングが拓く機械学習 と計算技術の新時代 (量子システム 推定の数理)

AUTHOR(S):

大関, 真之

CITATION:

大関, 真之. 量子アニーリングが拓く機械学習と計算技術の新時代 (量子システム推定の数理). 数理解析研究所講究録 2017, 2059: 13-23

ISSUE DATE:

2017-10

URL:

<http://hdl.handle.net/2433/237210>

RIGHT:

量子アニーリングが拓く 機械学習と計算技術の新時代

東北大学 大学院情報科学研究科応用情報科学専攻 * 大関 真之

Masayuki Ohzeki

Graduate School of Information Sciences, Tohoku University

§1. 前書き

量子アニーリングと呼ばれる最適化問題を解く汎用的解法がある。一般的に量子という言葉を目にする場面は研究者界隈でしかないだろうが、驚くべきことにこの名詞が Web メディアを始め、TV ニュースにまでなる日が来た。量子力学という非常に小さなスケールにおいて起きる常識離れたルールに基づく挙動を利用して、実際の社会問題の解決に結びつく最適化問題の効率的な解法、それが量子アニーリングである。この量子アニーリングは、もともと数値計算手法、アルゴリズムのひとつとして提案されたに過ぎない。しかしそのアイデアを元に、実際に量子力学に支配されて動作する系を利用して、全くそのままに動作する量子力学による計算機を実現にまで至ったためにメディアを巻き込んで世界中で騒然となっている。本稿では、その量子アニーリングの基本的な部分と応用的な部分について紹介していこう。

§2. 量子アニーリング

量子アニーリングの原理 [1] は至って単純である。解きたい最適化問題を磁性体の数理解模型であるイジング模型により記述するだけで実行可能である。特定のコスト関数を最小化したいという問題に対して、そのコスト関数を記述するイジング模型のハミルトニアン

を用意する.

$$\hat{H}_0 = - \sum_{i \neq j} J_{ij} \hat{\sigma}_i^z \hat{\sigma}_j^z - \sum_{i=1}^N h_i \hat{\sigma}_i^z. \quad (1)$$

ここで $\hat{\sigma}_i^z$ はパウリ行列の z 成分である. 相互作用の強さ J_{ij} と局所磁場の強さ h_i が解きたい最適化問題に応じて決定される. 添え字にある i はイジング模型の自由度であるスピンの配置された場所を示す. 各場所にはスピン $1/2$ の量子力学的状態が存在するようにシステムを用意する必要がある. 量子アニーリングを実装したとされるマシンやチップは, この量子力学的状態を超伝導状態を利用するなどして, 擬似的に解くべきイジング模型を用意することで実現している. このイジング模型が示すシステムのエネルギーを最小化するということを目指すことで最適化問題を解くというロジックである.

量子アニーリングではその解きたい最適化問題を解くために量子揺らぎを導入する. 導入する量子揺らぎの条件は, 解きたい最適化問題とは交換しない項を導入すること, また自明な基底状態を持つことが望まれる. これらの条件を満たす量子揺らぎとして, 頻繁に利用されるのが, 横磁場である. 横磁場の効果を示すハミルトニアンを以下のように用意する.

$$\hat{H}_1 = -\Gamma \sum_{i=1}^N \hat{\sigma}_i^x. \quad (2)$$

この横磁場の基底状態は, すべてのスピンの横向きになっている状態となる. スピン $1/2$ 状態では, $\hat{\sigma}^z$ の二つの固有状態からなるスピン上向きの状態と下向きの状態をとるが, これらの重ね合わせの状態を横向きの状態と呼ぶ.

量子アニーリングは, これら二つのハミルトニアンを時間変化を伴う係数により組み合わせたシステムを用意する.

$$\hat{H}(t) = f(t) \hat{H}_0 + (1 - f(t)) \hat{H}_1. \quad (3)$$

この時間変化を伴う係数は理論的な考察を行う際には, 簡単のため $f(t) = t/\tau$ として, 量子アニーリングを実行する時間 τ に対して, 規格化された時間 $s = t/\tau$ に対して線形に変化させることが多い. しかし実際には実装に関する制約のためや, 効率の良い基底状態の探索のために線形とは限らない. いずれにせよ, $f(0) = 0$ 及び $f(\tau) = 1$ として, 時間変化をするハミルトニアンを利用して, 量子状態を駆動させる. その結果, 最終時刻 τ において最適化問題を表すハミルトニアンの基底状態を持つ確率が高いことを期待するのが量子アニーリングである.

§3. 断熱定理と量子アニーリングの実際

量子アニーリングでは、時間依存をするハミルトニアンにより量子状態を駆動させる。その際に非常にゆっくりと時間発展をさせることで、理想的には断熱定理を利用する。断熱定理のあらすじを先に述べておくと、初期状態として初期時刻のハミルトニアンの基底状態を設定した上で、非常にゆっくりとした時間発展をさせると、各時刻におけるハミルトニアンの基底状態をたどることが可能である [2]。そのため、断熱定理の条件下で時間発展をさせれば、横磁場の基底状態から出発した量子状態は、所望の最適解に対応する量子状態へと変化していくことが可能となる。断熱定理そのものは形式的に部分積分を逐次的に実行すれば導出することができる。量子アニーリングの原理そのものはこれで全てを尽くしている。量子状態特有の難しいエンタングルメントといった概念は、量子アニーリングの基本的原理を理解するためには必要ではない。この事実が量子アニーリングを利用する場合に、参入障壁が低いということのできる所以である。

さてカナダのベンチャー企業である D-wave Systems 社のマシンについて少しここで述べることにしよう。量子アニーリングの原理に従い、超伝導量子ビットを操ることで最適化問題を解くことを目指したマシン、それが D-wave マシンだ。2017 年 5 月時点で 2000 量子ビットのイジング模型で記述される最適化問題の解を求めることができる。

さて、このマシンは断熱定理に従った振る舞いをしているのだろうか？答えは No である。残念ながら D-wave マシンは現状では断熱定理に従った正統な量子アニーリングマシンとは呼べない。というよりも、量子アニーリングの理論自体が、現実には即していないというべきだろうか。上記のシュレーディンガー方程式の考察においては、環境との相互作用について一切考慮していないのだ。横磁場と解きたい最適化問題を表すハミルトニアン他に、実際には環境との相互作用が現れるために断熱定理に従った想定した振る舞いとは全く異なることが起こることが指摘されている。断熱定理の教えによれば、非常にゆっくりと時間発展をすれば基底状態に到達する確率が増大することになる。しかしながら環境の相互作用を考慮すると、現状の量子アニーリングにかかる時間では、断熱定理がもはや破綻してしまい、ある有限温度の平衡状態へと到達することが言われている [3]。この性質は一見すると量子アニーリングが理論通りに実現していないことを示しており、残念なことに思うかもしれない。しかし D-wave Systems は機転を利かせた。有限温度の平衡状態に到達するということは、量子アニーリングを終えて得られる結果は、基底状態とは限らず、様々な状態が実現することになる。同じ条件で何度も実行すれば、 \hat{H}_0 で指定されるハミルトニアンによる平衡状態で生じるスピン配位のサンプリングを実行すること

になる。このサンプリングを実行するには、後述するマルコフ連鎖モンテカルロ法を利用することが多い。このマルコフ連鎖モンテカルロ法では、所望の平衡状態に到達させるまでにやや時間のかかる傾向があり、効率的なアルゴリズムの開発が重視されてきた。いわば量子アニーリングの失敗作が、この時間のかかるサンプリングの解決方策となり得るのだ。D-wave Systems 社はある時から、最適化問題を解くための量子アニーリングマシンという表現から、機械学習分野におけるサンプリングにも有効だという宣伝文句を付け足すようになった。その背景には、こういった量子アニーリングの失敗がある。

§4. 機械学習への応用

ここでやや量子アニーリングからは離れるが、サンプリングと呼ばれる手法が重要な役目を果たす機械学習の一つの応用例を紹介しよう。

§4.1. ボルツマン機械学習

計測技術や信号処理技術、そして情報処理そのものの質の向上により、我々は大量のデータを取得することが可能となった。 N 次元のベクトル $\mathbf{x}^{(d)}$ で表される大量のデータ ($d = 1, 2, \dots, D$) が与えられたときに、そのデータの起源を少数の説明変数で表した生成モデルを解明する処方箋のひとつがボルツマン機械学習である。ここで大前提として、データは確率的に出力されて得られているものとする。

ボルツマン機械学習では、統計力学で基本となるカノニカル分布に従ってデータが出力されると考える。

$$P(\mathbf{x}|\mathbf{u}) = \frac{1}{Z(\mathbf{u})} \exp \{-E(\mathbf{x}|\mathbf{u})\}. \quad (4)$$

ここで $Z(\mathbf{u})$ は分配関数、 \mathbf{u} はデータの構造を表すエネルギー関数 $E(\mathbf{x}|\mathbf{u})$ を形作るパラメータである。例えば N 次元の 2 値データがイジング模型のカノニカル分布から生成されたと仮定する場合には、

$$E(\mathbf{x}|\mathbf{u}) = -\sum_{i \neq j} J_{ij} x_i x_j - \sum_{i=1}^N h_i x_i \quad (5)$$

とする。 $x_i = \pm 1$ であり、パラメータ $\mathbf{u} = (J, \mathbf{h})$ によりエネルギー関数が特徴づけられている。上記の仮定のもと、与えられた大量のデータの経験分布

$$P_D(\mathbf{x}) = \frac{1}{D} \sum_{d=1}^D \delta(\mathbf{x} - \mathbf{x}^{(d)}) \quad (6)$$

に最も近いカノニカル分布を探してることがボルツマン機械学習の目標となる。その結果、パラメータ \mathbf{u} によりデータの経験分布を「もっともらしく」再現する確率分布を得ることができる。パラメータそのものからデータの特徴を調べることも可能である。

さてそうすると2つの異なる確率分布を持ってきたときに、それらが近いか遠いかを調べるための計量が必要だ。最も一般的に用いられるのがカルバック・ライブラー (KL) 情報量である。

$$D_{\text{KL}}(P|Q) = \int d\mathbf{x} P(\mathbf{x}) \log \left(\frac{P(\mathbf{x})}{Q(\mathbf{x})} \right). \quad (7)$$

この KL 情報量の意味で、データの経験分布に最も近い確率分布を与えるパラメータ \mathbf{u} を求めてみよう。 $Q(\mathbf{x})$ を未知のパラメータを持つ $P(\mathbf{x}|\mathbf{u})$ に、 $P(\mathbf{x})$ をデータの経験分布 $P_D(\mathbf{x})$ としよう。このとき KL 情報量の最小化問題は、以下の最大化問題と等価であることが分かる。

$$\mathbf{u}^* = \arg \max_{\mathbf{u}} L(\mathbf{u}). \quad (8)$$

ここで $L(\mathbf{u})$ は対数尤度関数 (の経験平均) と呼び、以下のように定義される。

$$L(\mathbf{u}) = \frac{1}{D} \sum_{d=1}^D \log P(\mathbf{x} = \mathbf{x}^{(d)} | \mathbf{u}). \quad (9)$$

ボルツマン機械学習は、この対数尤度関数の最大化 (最尤法) を行うことで、得られたデータに適合する「もっともらしい」パラメータの推定を行うとも言換えることができる。

最尤法を実行するためには、対数尤度関数の微分を逐次的に足していく勾配法を利用する。

$$\mathbf{u}[t+1] = \mathbf{u}[t] + \eta \frac{\partial L(\mathbf{u})}{\partial \mathbf{u}}. \quad (10)$$

ここで η は学習係数と呼ばれる量で、小さければ小さいほど正確であるが計算時間の長大化に繋がるのでほどよい値をとることが要求される。対数尤度関数の微分が必要となるので、パラメータ \mathbf{u} について対数尤度関数の微分を取ってみる。

$$\frac{\partial L(\mathbf{u})}{\partial \mathbf{u}} = -\frac{1}{D} \sum_{d=1}^D \frac{\partial E(\mathbf{x} = \mathbf{x}^{(d)} | \mathbf{u})}{\partial \mathbf{u}} + \left\langle \frac{\partial E(\mathbf{x} | \mathbf{u})}{\partial \mathbf{u}} \right\rangle_{\mathbf{u}}. \quad (11)$$

第1項はエネルギー関数の形を知っていれば評価は容易である。データに関する経験平均をとるだけだ。一方第2項は熱平均の計算 $\langle \cdots \rangle_{\mathbf{u}} = \sum_{\mathbf{x}} \cdots \times P(\mathbf{x} | \mathbf{u})$ が必要となる。そこでエネルギー関数 $E(\mathbf{x} | \mathbf{u})$ で指定された平衡状態を実現するシミュレーションを行うことで、その熱平均を計算することにする。そのシミュレーション法として採用されるのがマルコフ連鎖モンテカルロ法である。このマルコフ連鎖モンテカルロ法を実行すること

で、ある特定の平衡状態を模した確率分布を生成することができる。その確率分布に従ったスピン配位をサンプリングすることで期待値を計算する。

このマルコフ連鎖モンテカルロ法によるサンプリングの実施の代わりに、D-wave マシンを利用するというわけだ。量子アニーリングマシンとしての失敗にくじけることなく、前向きに利用方法を提案する精神には舌を巻く思いだ。さらに D-wave Systems 社も黙っていない。断熱定理が有効となる条件に当てはまるように技術革新を繰り返して、真の量子アニーリングマシンの実現に向けて努力をしている。目的の達成に向けて努力をし続ける中の、あくまでスピンアウト作品として、このボルツマン機械学習への応用がある。

他にも機械学習分野におけるアンサンブル学習や強化学習、さらには辞書学習など、最適化問題を通じて実行される各種アルゴリズムへの利用も提案されている。

§4.2. アンサンブル学習

基本的に機械学習では、入力 \mathbf{x} に対して出力 y を返す非自明な関数を獲得することを目標としている。どのような関数の中で探すか、により方法が分かれる。探す関数のクラスを限定して、できるだけ近いものを探すというのが目標となる。できるだけ近いものを探すという部分に量子アニーリングを利用しようという方針だ。

その代表例とされるのが、QBoost と呼ばれる量子アニーリングを利用したアンサンブル学習への適用だ [4]。アンサンブル学習では「三人寄れば文殊の知恵」という格言の通り、複数の識別器を利用することで高性能な識別精度を引き出すこと（ブースティング）を目的とする。適当に用意された識別器を $c_i(\mathbf{x}) \in \{-1, 1\}$ とする。ここで \mathbf{x} は識別される対象となるデータを表す。この識別器自体は大した性能を持たないとする。その性質から弱識別器と呼ばれる。これらを適当な重みをつけて組み合わせた識別器を以下のように用意する。

$$C(\mathbf{x}) = \text{sign} \left(\sum_{i=1}^N w_i c_i(\mathbf{x}) \right). \quad (12)$$

ここで w_i を重みとして、二値 $w_i \in \{0, 1\}$ を取るとする。対応する弱識別器を使うか使わないかを選択する重みとなる。

この識別器に対して、 \mathbf{x} のラベル $y \in \{-1, 1\}$ を正解として与えることにより教師あり学習を行う。その際に以下の最適化問題を解くことを考える。

$$\min_{\mathbf{w}} \left\{ \sum_{d=1}^D \left(C(\mathbf{x}^{(d)}) - y^{(d)} \right)^2 + \lambda \sum_{i=1}^N w_i \right\}. \quad (13)$$

ここで第二項は、与えられたデータに対して過剰に適合しないようにするための正則化項

を表す。この最適化問題を素朴に式変形すると、 w_i の 1 次の項と 2 次の項のみが現れる。ここで $w_i = 2\sigma_i - 1$ ($\sigma_i \in \{-1, 1\}$) とすることで、イジング模型のハミルトニアンに対応する形が得られる。その結果、相互作用係数は、

$$J_{ij} = -\frac{1}{2} \sum_{d=1}^D c_i(\mathbf{x}^{(d)}) c_j(\mathbf{x}^{(d)}) \quad (14)$$

となり、局所磁場は

$$h_i = -\frac{\lambda}{2} + \sum_{d=1}^D c_i(\mathbf{x}^{(d)}) y^{(d)} - \frac{1}{2} \sum_{d=1}^D \sum_{j=1}^N c_i(\mathbf{x}^{(d)}) c_j(\mathbf{x}^{(d)}) \quad (15)$$

となる。この手続きを通して D-wave マシンにより最適化された識別器を獲得することができる。これにより画像から森林部分かどうかの診断に利用されたりと実用例も出て注目を集めている [5]。

§4.3. 強化学習

機械学習分野において一躍注目を集めているのは強化学習の進展であろう。プロ棋士をも凌駕する実力を有したアルファ碁の成長に伴い注目を集めている。その強化学習に現れる最適化問題において、D-wave マシンを利用した事例が報告されている。

ロボットなどの動的なシステムにおいて、意思決定を確率的に行うプロセスを素朴に定式化した一例としてマルコフ決定過程がある。現在の状況 s 、行動 a 、それに対してどんな行動を取るかを定めるポリシーとして π からなる 3 つの要素 (π, s, a) が確率的に変動する中、以下の Q 関数と呼ばれる報酬の期待値を最大化するようなポリシー π を選択することを考える。

$$Q(\pi, s, a) = \langle r(s, a) \rangle + \left\langle \sum_{t=1}^{\infty} \gamma^t r(\Pi_t^s, \pi(\Pi_t^s)) \right\rangle. \quad (16)$$

ここで $r(s, a)$ は即時報酬を表す。 $\gamma \in (0, 1)$ は報酬の減衰率である。さらに Π_t^s は時刻 i までの状態 s となったマルコフ過程の履歴を表しており条件付き確率 $P(s'|s, a)$ で生成される。 $\langle \dots \rangle$ は実現した状態、行動についての経験平均を取ることに相当する。この時、ポリシー π について最大化された $Q^*(s, a) = \arg \max_{\pi} Q(\pi, s, a)$ は、以下のベルマン方程式を満たすことが知られている。

$$Q^*(s, a) = \langle r(s, a) \rangle + \gamma \sum_{s'} P(s'|s, a) \max_{a'} Q^*(s', a'). \quad (17)$$

$Q^*(s, a)$ はこの方程式の自己無撞着解になる．そこで逐次繰返しをして収束して得られた解を利用することにする．ここで適当な初期条件 $Q_0(s, a)$ に対して、ベルマン方程式に n 回代入を繰返したものを $Q_{n+1}(s, a)$ と置くと、以下の Temporal Difference(TD) を計算することができる．

$$Q_{n+1}(s, a) - Q_n(s, a) = \langle r(s, a) \rangle + \gamma \sum_{s'} P(s'|s, a) \max_{a'} Q_n^*(s', a') - Q_n(s, a). \quad (18)$$

この TD に基づく強化学習の方法を Q 学習と呼ぶ．TD が 0 であれば $Q(s, a)$ の収束解が得られたものと解釈できる．そこで TD を $Q(s, a)$ のある種の勾配と捉えて学習を進める．この TD の計算には条件付き確率による期待値計算を含むため、マルコフ連鎖モンテカルロ法を実行する必要がある計算時間の長大化が問題となる．そこで $Q(s, a)$ をイジング模型の自由エネルギーであると捉えることにより、以下のように近似することにする．

$$Q(s, a) \approx -F(s, a). \quad (19)$$

ここでどんなイジング模型を対応させるかに任意性があるが、隠れ変数ありの制限ボルツマンマシンを利用した先行研究 [6] に倣い、状態 s と行動 a の変数、さらに隠れ変数としてイジング変数を余分に用意して、これらの間を結合させたイジング模型を用意する．具体的には

$$\begin{aligned} -F(s, a) = & \sum_i \sum_j w_{ij} s_i \langle \sigma_j \rangle + \sum_i \sum_k w_{ik} a_i \langle \sigma_k \rangle \\ & + \sum_{i,i'} u_{ii'} \langle \sigma_i \sigma_{i'} \rangle - \frac{1}{\beta} \sum_{\sigma} P(\sigma|s, a) \log P(\sigma|s, a). \end{aligned} \quad (20)$$

という形を保つイジング模型が対応する．前の 3 項がハミルトニアン¹の熱平均に対応しており、最後の項がエントロピー項に対応する． w_{ij} や w_{ik} がそれぞれ状態 s_i と行動 a_i とイジングスピン h_j などとの相互作用を表しており、 $u_{ii'}$ がイジングスピン同士の相互作用を表す．隠れ変数部分として導入したものはイジング模型であるから、D-wave マシンなどでイジング型計算処理を行うことで、高速にサンプリングを行うことができるため、期待値の計算が容易である．ひいては自由エネルギーの計算も非常に高速に行うことができることになり Q 学習にとって最大の問題であった部分が解消される．実際 D-wave 等にイジング模型を用いたソフトウェアを提供する 1Qbit のメンバーにより、上記の Q 学習に対する量子アニーリングマシンの利用が提案されて、その検証が進んでいる [7]．

§5. 確率解釈不能系の進展

§5.1. 鈴木トロッター分解

量子アニーリングの基本的なスキームでは、解きたい最適化問題を表したハミルトニアン \hat{H}_0 に対して、横磁場をかけた模型を考える。ここでパウリ行列の交換関係より、 z 成分と x 成分では交換しないことから、鈴木トロッター分解により、以下のような近似を考える。

$$Z = \text{Tr} \left\{ \prod_{t=1}^{\tau} \langle \sigma_{t+1} | \exp \left(-\frac{\beta}{\tau} \hat{H}_0 \right) \exp \left(\frac{\beta\Gamma}{\tau} \sum_{i=1}^N \hat{\sigma}_{it}^x \right) | \sigma_t \rangle \right\} + O \left(\frac{1}{\tau^2} \right). \quad (21)$$

ここで $|\sigma_t\rangle$ はパウリ行列の z 成分を対角化する表示における固有ベクトルである。 $\hat{\sigma}^x$ に関係するところを以下のように恒等式を用いて書き直す。

$$\exp \left(\frac{\beta\Gamma}{\tau} \hat{\sigma}_{it}^x \right) = \cosh \left(\frac{\beta\Gamma}{\tau} \right) + \hat{\sigma}_{it}^x \sinh \left(\frac{\beta\Gamma}{\tau} \right). \quad (22)$$

さらに $\exp(-2\gamma) = \tanh(\beta\Gamma/\tau)$ を用いて、

$$\exp \left(\frac{\beta\Gamma}{\tau} \hat{\sigma}_{it}^x \right) = \cosh \left(\frac{\beta\Gamma}{\tau} \right) \exp(-\gamma) \begin{pmatrix} \exp(\gamma) & \exp(-\gamma) \\ \exp(-\gamma) & \exp(\gamma) \end{pmatrix} \quad (23)$$

という表示に変える。この行列はパウリ行列の z 成分を使って、 $\exp(\gamma\sigma_{it}\sigma_{it+1})$ として表現することができる。その結果、横磁場を持つスピン模型は、以下のようにトロッター方向に同じ古典ハミルトニアンを持ち、トロッター方向間に強磁性相互作用を持つ模型に変換される。

$$Z = \left(\cosh \left(\frac{\beta\Gamma}{\tau} \right) \exp(-\gamma) \right)^{N\tau} \sum_{\sigma} \exp \left(-\frac{\beta}{\tau} \sum_{t=1}^{\tau} \hat{H}_0 + \gamma \sum_{i=1}^N \sum_{t=1}^{\tau} \hat{\sigma}_{it} \hat{\sigma}_{it+1} \right) + O \left(\frac{\beta^2}{\tau} \right). \quad (24)$$

変換が厳密に行えるのは、トロッター数 τ が無限に大きい場合である。意味のある結果を取り出すためには逆温度 β も大きくなければならない。そこで標準的には $\beta/\tau = 1$ として、 τ と β を同程度に大きくする。

§5.2. 確率解釈不能なハミルトニアン

標準的な量子アニーリングでは鈴木トロッター分解を行うことにより実効的にパウリ行列の z 成分だけの模型、古典系で書くことができる。その書き換えは分配関数に現れる和

を取られる項に関して等価な寄与をするものを書き換えている。この部分はボルツマン因子と呼ばれ、あるスピン配位が実現する確率を表す部分である。量子系の分配関数を記述するボルツマン因子を、鈴木トロッター分解などを行い、確率として解釈ができる場合、その対象としている物理系のことを Stochastic 系とよぶ。一方でボルツマン因子が負となる場合、負符号が生じると言われる。例えば量子アニーリングで横磁場以外の量子揺らぎを導入した場合には負符号が生じる。このような場合、確率としてボルツマン因子を解釈することが難しいため、Non-stochastic 系と呼ぶ。Stochastic 系は量子系であっても対応する古典系が存在するため、量子系であっても確率的シミュレーションを行うことが可能である。一方で Non-stochastic 系は、対応する古典系には負符号が生じるなど不都合な点があるため効率的に確率的シミュレーションを行うことができないため、古典的な世界と量子的な世界の境目が見え隠れしている重要な問題設定であることがうかがえる。

最近この Non-stochastic 系が急速に注目を集めている。きっかけは量子アニーリングに横磁場以外の量子揺らぎを導入することによって指数関数的加速を示した例 [8, 9] があることと、上記のように対応する古典系がある量子系では、本当の量子性の有効な利用をしたとは言えないことに人々が気づいたことにある。前者の結果については当初量子性の表す非自明な加速や最適化問題を解く性能の飛躍的向上を示す好例ではないかと考えられたが、最近筆者による適応的量子モンテカルロ法 [10] の発見により、素朴には負符号が生じるような問題であっても適切に変換を施すことによって負符号を回避した古典系が存在することが示された。しかしながら単純な古典系ではなく、横磁場以外の量子揺らぎの存在により、常に横磁場が変化するような特殊な系であることが示された。そのため単純な横磁場とは異なる振る舞いをする可能性がある。また D-wave Systems を始め、量子アニーリングマシンを追いかけて開発を進める各研究機関では、横磁場以外の量子揺らぎを利用した量子アニーリングの実装を目指している。

実は横磁場以外の量子揺らぎを導入することにより、量子アニーリングは、いわゆる従来知られている量子コンピュータの形式であるゲート方式が目指してきた万能量子コンピュータに昇格することが知られている [11]。ゲート方式は常に生じる誤りに対してフォールトトレラントなシステムを構築するために、その実現に時間がかかっている。一方で量子アニーリング方式は、エネルギーの低い状態を利用するために比較的安定したシステムとなっているために、その実現に時間がかからなかったのだ。そのため、量子アニーリング方式を推し進めて人々の夢である万能量子コンピュータを実現させようという動きがある。その鍵を握るのが Non-stochastic 系の理解と実現にある。

謝辞

本研究は科学研究費補助金 基盤研究 B「量子アニーリングが拓く機械学習と計算技術の新時代」の支援を受けて実施されている。

参考文献

- [1] T. Kadowaki and H. Nishimori, Phys. Rev. E 58, 5355 (1998).
- [2] S. Suzuki and M. Okada, Journal of the Physical Society of Japan 74, 1649 (2005).
- [3] M. H. Amin: Physical Review A, 92, 052323 (2015).
- [4] H. Neven and V. S. Denchev and G. Rose and W.G. Macready: ACML, 333 (2012).
- [5] E. Boyda, S. Basu, S. Ganguly, A. Michaelis, S. Mukhopadhyay and R. R. Nemani, PLOS ONE, 12, 1 (2017).
- [6] B. Sallans and G. E. Hinton: JMLR, 1063 (2004).
- [7] D. Crawford, A. Levit, N. Ghadermarzy, J.S. Oberoi and P. Ronagh: arXiv preprint arXiv:1612.05695
- [8] Y. Seki and H. Nishimori: Physical Review E 85, 051112 (2012).
- [9] Y. Seki and H. Nishimori: Journal of Physics A: Math. and Theor., 48, 335301 (2015).
- [10] M. Ohzeki: Scientific Reports, 7, 41186 (2017).
- [11] J. D. Biamonte and P. J. Love: Physical Review A. 78, 012352 (2008).